Preservation Planning and PREMIS

Lisa Gregory

State Library of North Carolina
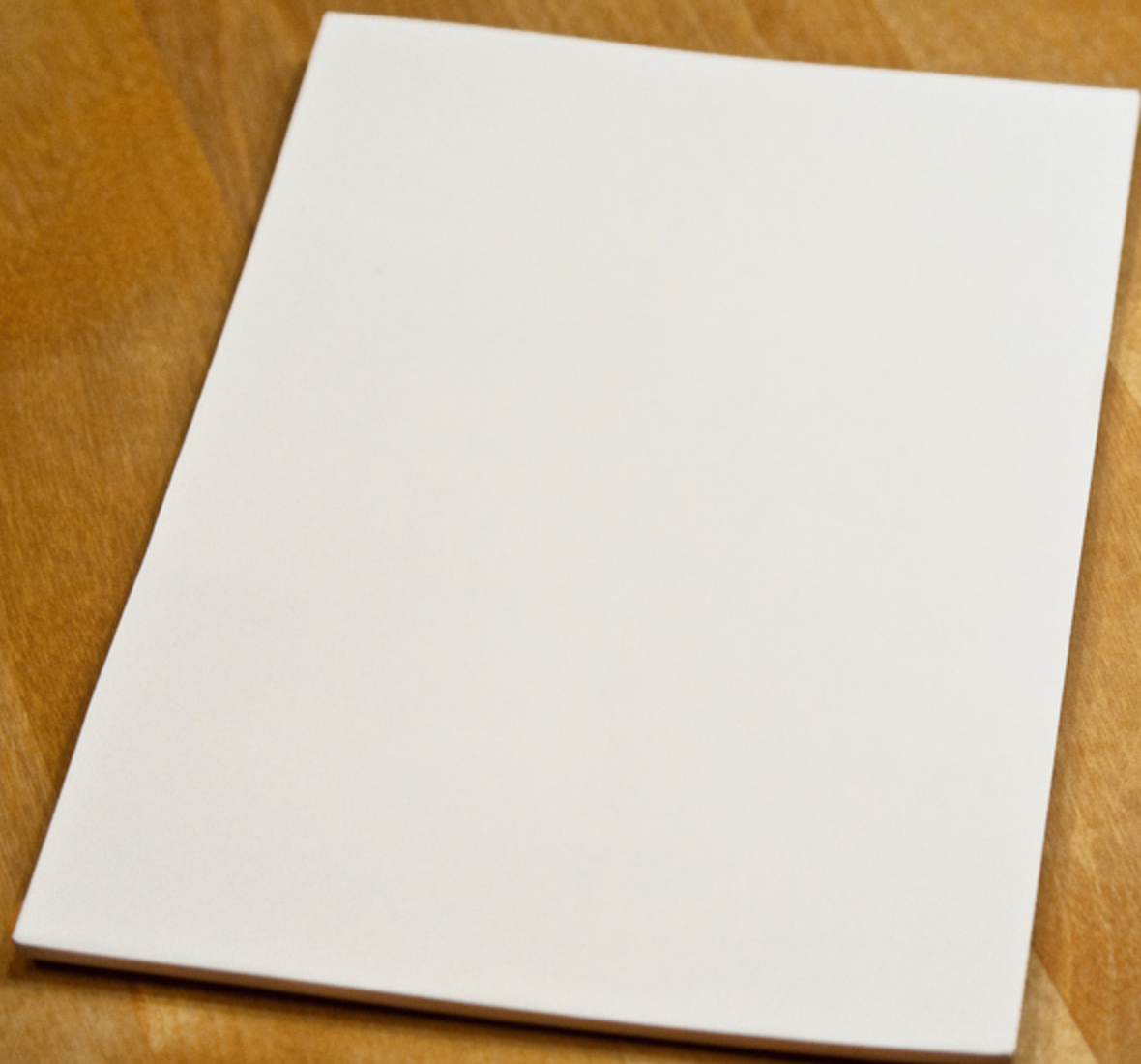
April 2013

Handout (URLs)

http://1.usa.gov/ZwNYKx

> Then, see top of page
for handout

INDIVIDUAL WORKSTATIONS

EXTERNAL DRIVES   REMOVABLE

SHARED LOCAL STORAGE   MEDIA

LAB/SHARED   INSTITUTIONAL

WORKSTATIONS   BACKUPS

MOBILE   CONTENT MANAGEMENT

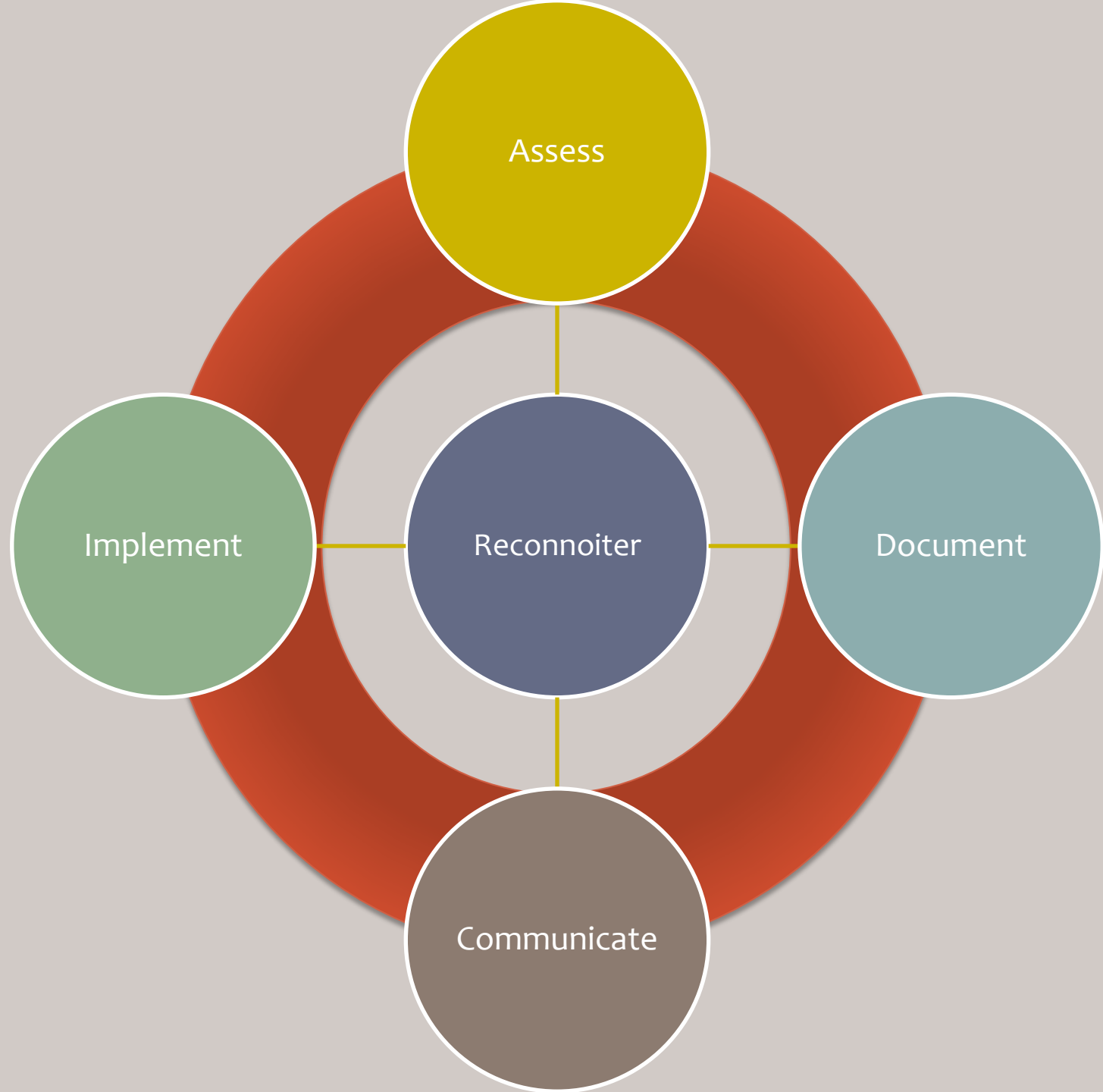DEVICES   SYSTEMS

INTERNAL/EXTERNAL SERVICES

HOT
MESS

Assess

Document

Communicate

Implement

Reconnoiter

Reconnoiter

Reconnoiter
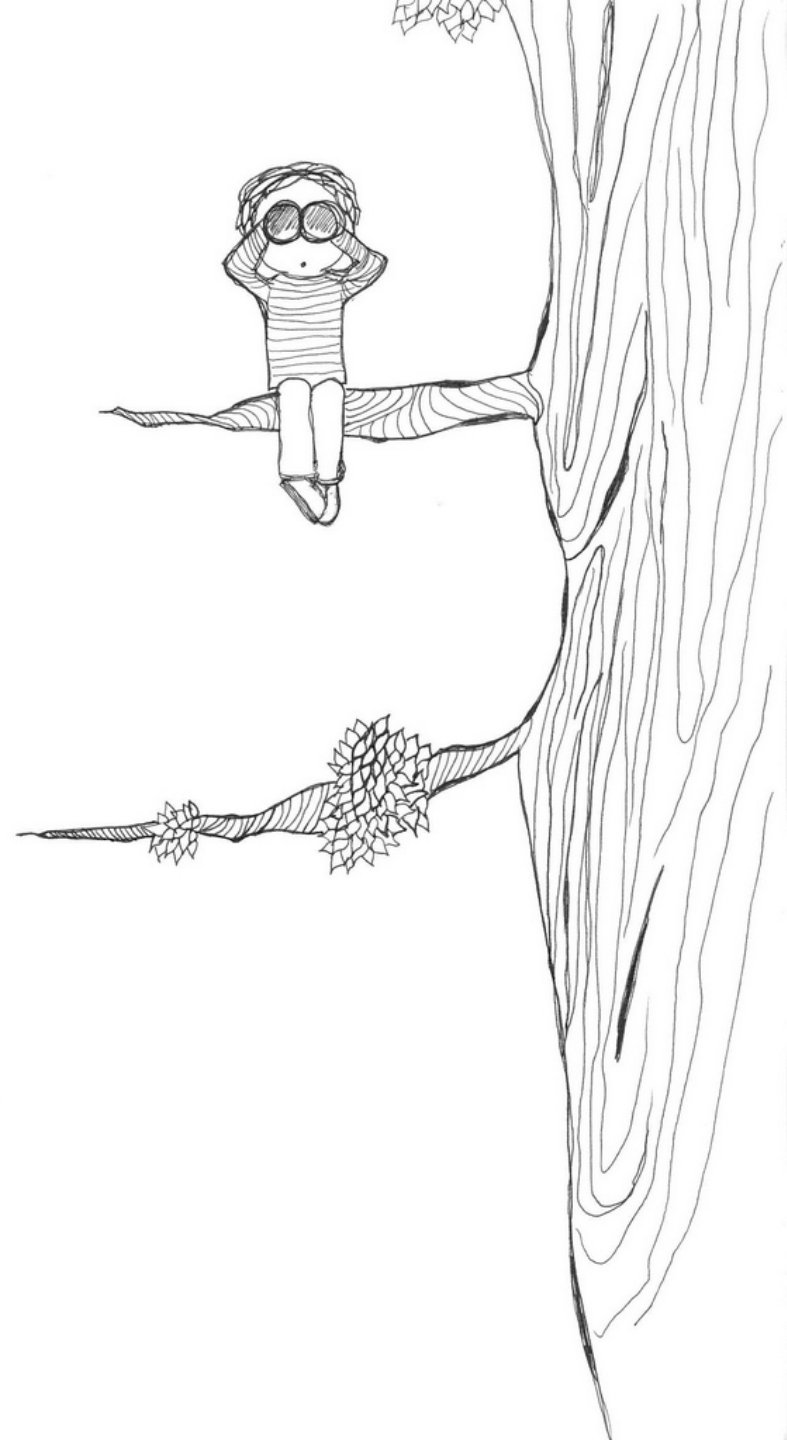
- The Signal:
  http://blogs.loc.gov/digitalpreservation/
- Twitter: search on ["digital preservation" or "digitalpreservation"] and start following those who post
- D-Lib Magazine

Assess

Assess

- Preservation statements: http://www.dlib.org/dlib/january13/webb/01webb.html

- SCORE: http://scoremodel.org/en

- TRAC: http://www.crl.edu/archiving-preservation/digital-archives/metrics-assessing-and-certifying/trac

Assess: TRAC/TDR

Organizational Infrastructure

Example: "A3.2 Repository has procedures and policies in place, and mechanisms for their review, update, and development as the repository grows and as technology and community practice evolve. "

Digital Object Management

Example: "B2.5 Repository has and uses a naming convention that generates visible, persistent, unique identifiers for all archived objects (i.e., AIPs)."

Technologies, Technical Infrastructure & Security

Example: "C1.2 Repository ensures that it has adequate hardware and software support for backup functionality sufficient for the repository's services and for the data held, e.g., metadata associated with access controls, repository main content."

IF YOU DONT WRITE IT DOWN IT NEVER HAPPENED

Document

Document

- Write down…
  - What you're promising to stakeholders
  - Results of assessment
  - Workflows

- Be transparent

- Bring in as many involved parties as is feasible

- Yes, you'll need to review it

Document

Preservation Policy Template

Digital Preservation Policy & Planning Workshop

MetaArchive
COOPERATIVE

2010-10-15

Policy Statement
This is a simple statement that relates digital preservation to the institution's mission and the communities it serves.

Summary Statement (try to keep to 1/2 page total)
This is a set of simple paragraphs that summarize the overall intent of the institution. May include the following elements:
- What are digital resources?
- Who makes selection decisions?
- When should decisions regarding preservation be made?
- What forms are prioritized (masters vs use copies)?
- Where does metadata fit?
- Why does it preserve content (e.g., institutional, legal, consortial obligations)?
- Who wrote this policy?
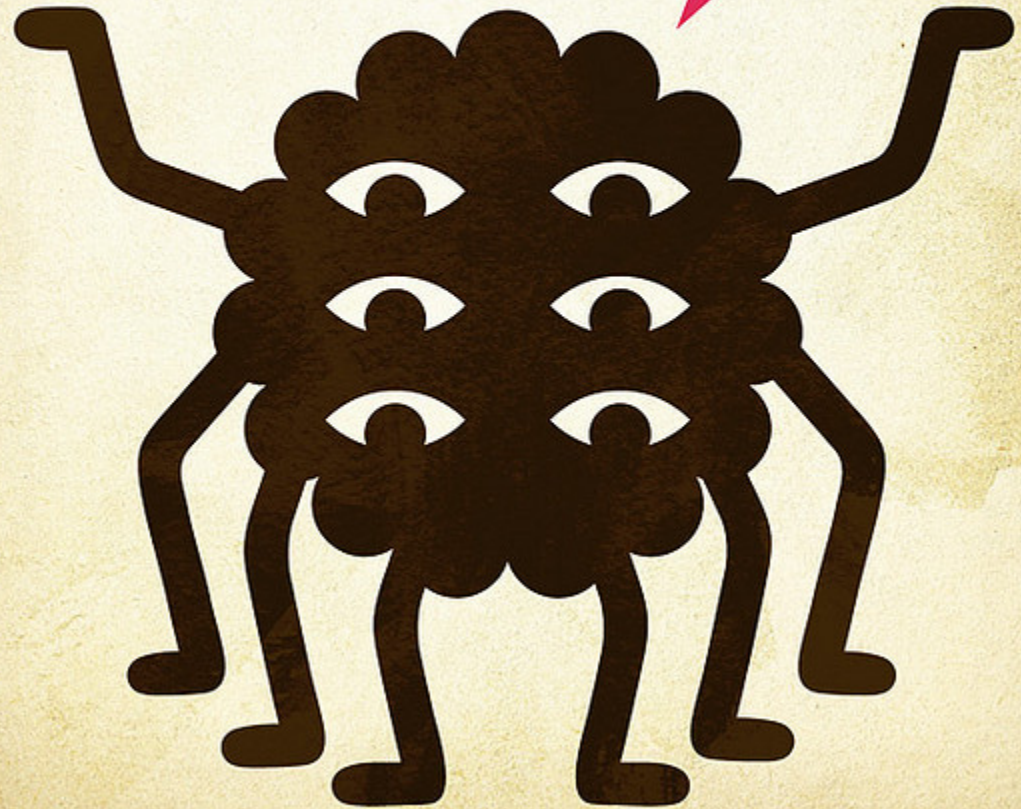- How often is this policy re-evaluated and by whom?

Scope
This section summarizes the resource groups (e.g., units, departments, or external parties) for which the institution takes responsibility and prioritizes these according to institutional importance.

Selection Criteria
This section outlines the way decisions are made regarding what will be preserved. May include the following:
- Mission
- Importance to user communities
- Quality of the digital resource
- Risk to the items/collections

http://metaarchive.org/public/resources/pres_comm/policy_planning/Digital_Preservation_Policy_Template.pdf

Communicate

Implement

Implement: Storage

- You know where stuff is, now…
  - Is what you're taking in matching your resources?
  - Is it organized according to a system you've defined?
  - Who can access/write to what parts?

Implement: Obsolete Media

- ## You know what you have, now…
  - Do you have the staff expertise and equipment to extract/work with the data?
  - Can you at least determine if the content is viable/relevant to your collecting scope?
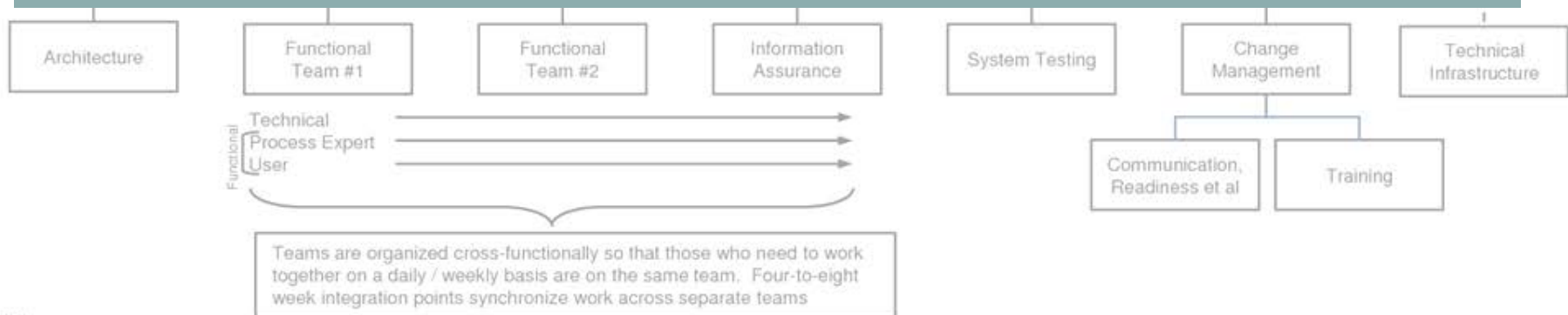 - Have you articulated what you will promise about such media going forward?

Implement: Organizational Infrastructure

Executive Sponsor

- ## You've written your policies, now...
  - How do you position yourself for administrative buy-in?
  - What could you reasonably improve within the next 3-5 years?
  - Are there other stakeholders headed your way?

| Architecture | Functional Team #1 | Functional Team #2 | Information Assurance | System Testing | Change Management | Technical Infrastructure |

Functional
- Technical
- Process Expert
- User

Communication, Readiness et al | Training

Teams are organized cross-functionally so that those who need to work together on a daily / weekly basis are on the same team. Four-to-eight week integration points synchronize work across separate teams

Implement : Workflows

- You've figured out what you *currently* do, now...
  - What small improvements could be made?
  - Are there tools that could help diminish the number of times a file is handled?
  - Can you start to think of an ideal workflow to work toward?

Implement: Metadata

- You've got your files, now...
  - What can you do to start recording/improve your preservation metadata?
  - Have you created a data dictionary or guidance document for that metadata?
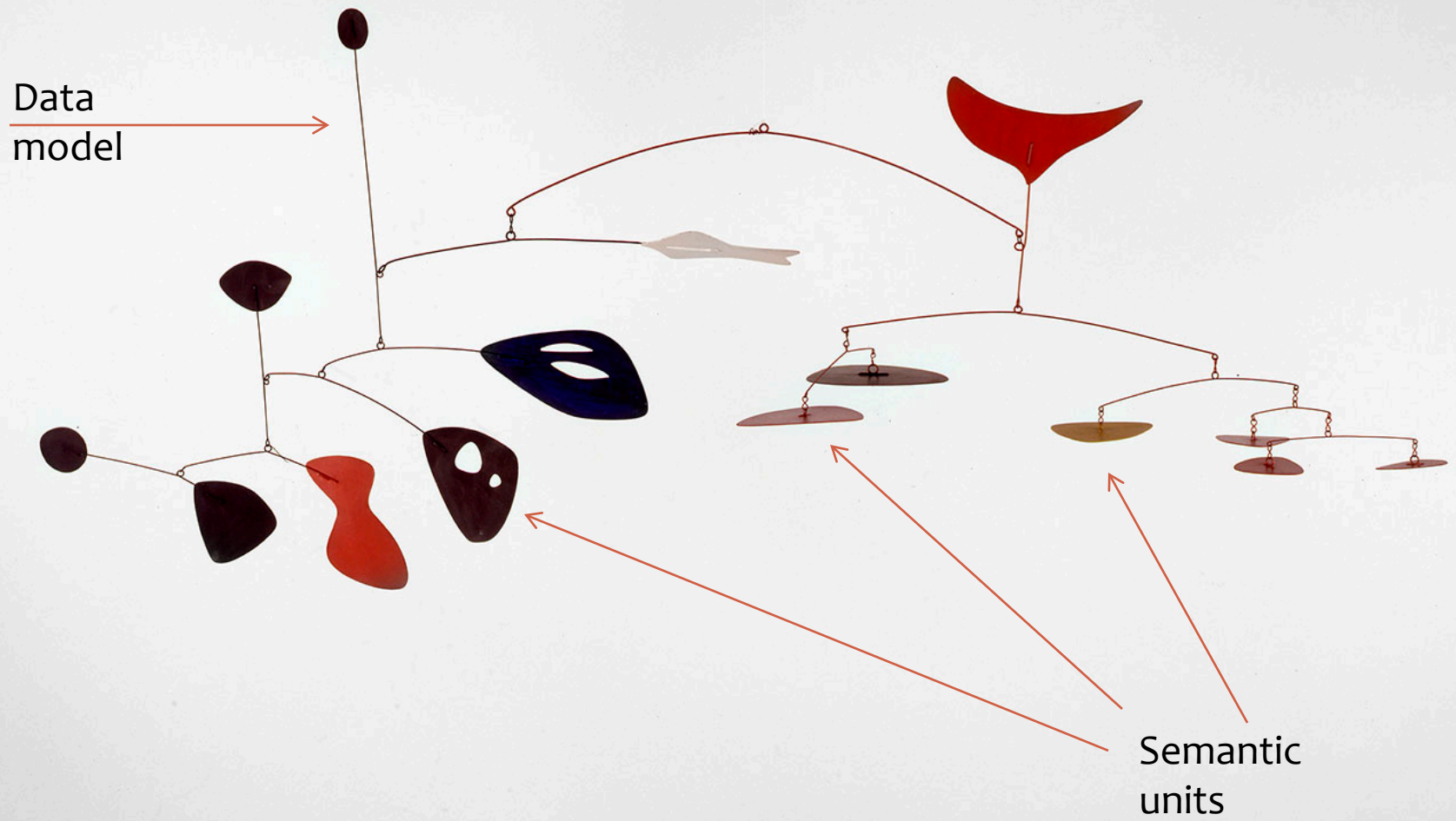 - Do you have audit trails for what happens to your files?

PREMIS

Data Model

- 5 entities
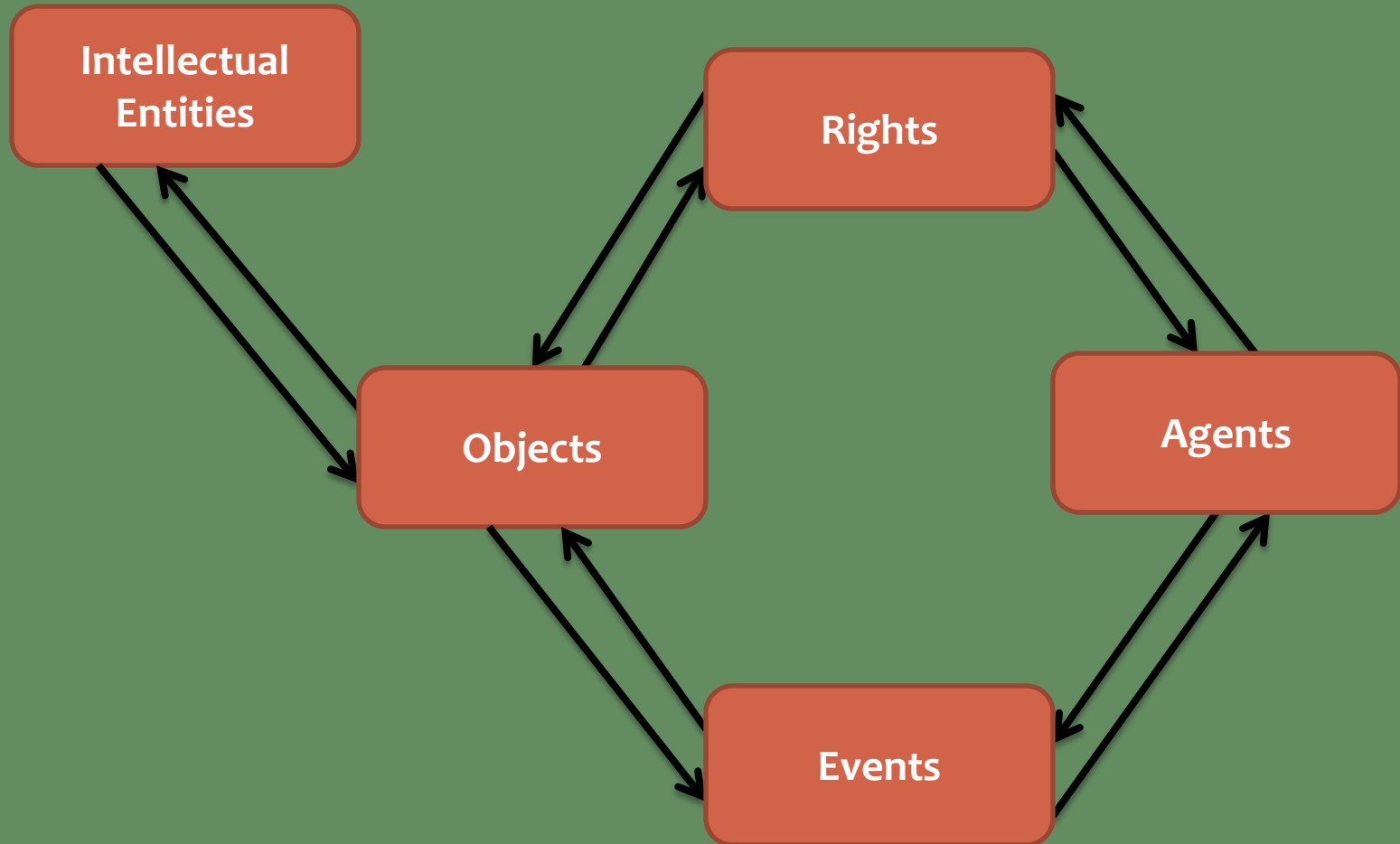- Organizes semantic units
- Defines relationships

- Defines semantic units
- Describes constraints for use
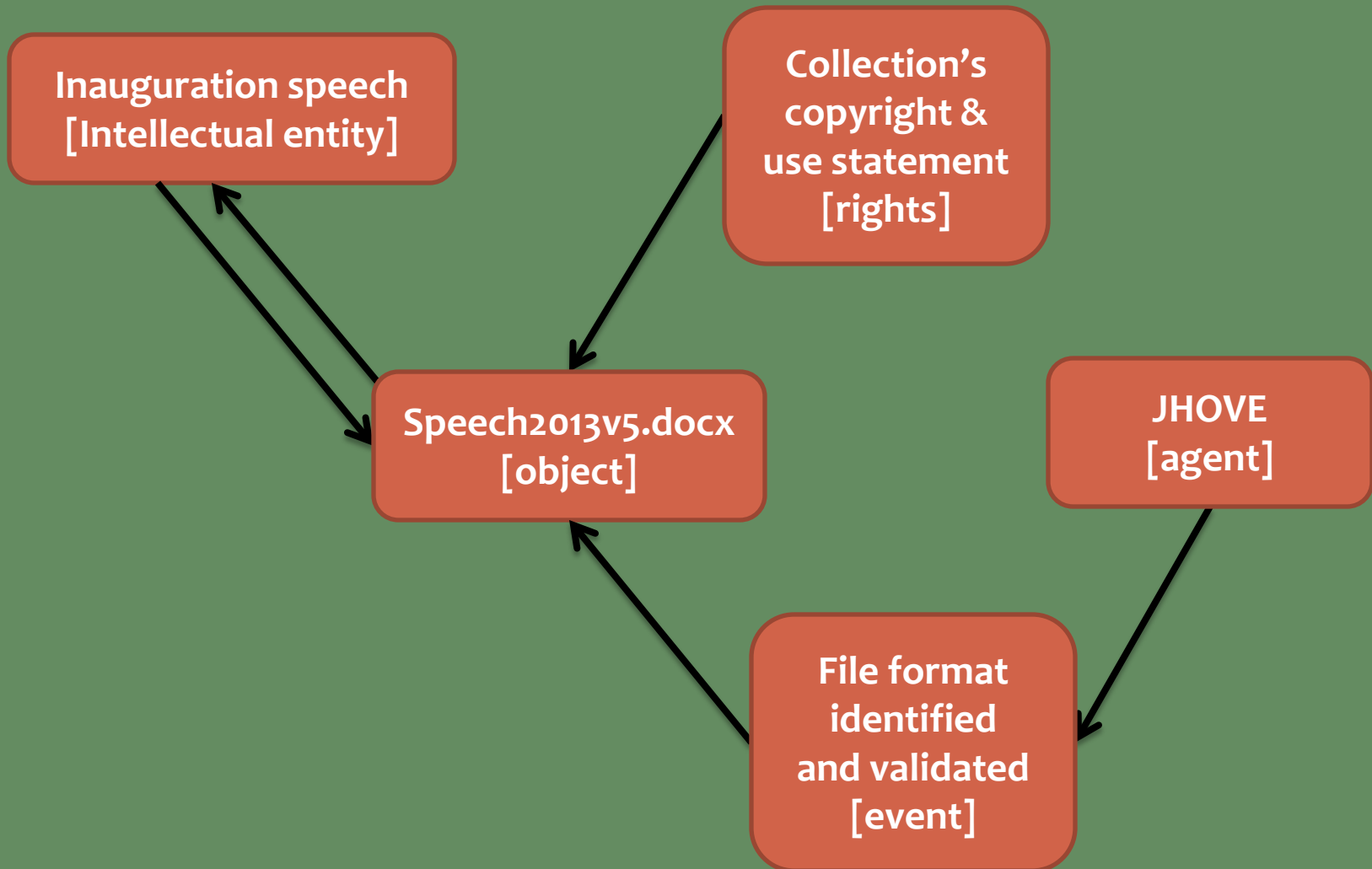- Gives examples

Data Dictionary

Data
model

Semantic
units

Alexander Calder, *Hanging Mobile*, 1951

Data Model: Example

PREMIS Data Dictionary: Main Semantic Units

1. Object Identifier
2. Object Category
3. Preservation Level
4. Significant Properties
5. Object Characteristics
6. Original Name
7. Storage
8. Environment
9. Signature Information
10. Relationship
11. Linking Event Identifier
12. Linking Intellectual Entity Identifier
13. Linking Rights Statement Identifier

PREMIS Data Dictionary: Example

Semantic unit	1.6 originalName		
Semantic components	None		
Definition	The name of the object as submitted to or harvested by the repository, before any renaming by the repository.		
Rationale	The name used within the preservation repository may not be known outside of the repository. A depositor might need to request a file by its original name. Also, the repository may need to reconstruct internal links for dissemination.		
Data constraint	None		
Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Not applicable
Examples		N419.pdf	
Repeatability	Not repeatable	Not repeatable	
Obligation	Optional	Optional	
Creation / Maintenance notes	This value would always be supplied to the repository by the submitter or harvesting application. How much of the file path to preserve would be up to the repository.		
Usage notes	This is the name of the object as designated in the Submission Information Package (SIP). The object may have other names in different contexts. When two repositories are exchanging content, it would be important for the receiving repository to know and record the name of the representation at the originating repository. In the case of representations, this may be a directory name.		

PREMIS Data Dictionary: Object Categories

Object category	Representation	File	Bitstream
Applicability	Applicable	Applicable	Not applicable
Examples		N419.pdf	
Repeatability	Not repeatable	Not repeatable	
Obligation	Optional	Optional	

North Carolina Digital Collections Preservation Metadata Implementation Guidelines

1. Bit Depth
2. Capture Tools
3. Checksum
4. Digital Characteristics (Access)
5. Digital Creation Date
6. Original Object Identifier
7. Preservation File Name
8. Preservation File Path
9. Resolution
10. Revision
11. Security
12. URL (Access)

North Carolina Digital Collections Preservation Metadata Implementation Guidelines

| Original Object Identifier-M	
Element	N/A
Mandatory	Yes
Repeatable	Yes
Crosswalk to PREMIS	1.6 originalName
Controlled vocabulary	None
Definition	An identifier for the master copy generated by the creator/submitter and formatted as received.
Input guidelines	1. Enter filename of the digital object with extension.
Examples	P1080067.JPG pubs_lawsofnorthcarol1817nort.pdf

Lisa Gregory
Digital Collections Manager
State Library of North Carolina
lisa.gregory@ncdcr.gov
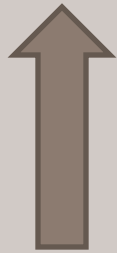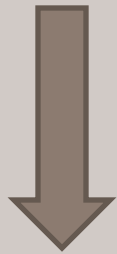
Credits

Me

Webinar Feedback

http://www.surveymonkey.com/s/2013_DP1